CONVOLUTIONAL NEURAL NETWORKS

BY SUNEESH JACOB

CONVOLUTIONAL NEURAL NETWORK

- Convolution
  - Kernel (Filter)
  - Stride
  - Padding
    - Zero
    - Same

CONVOLUTIONAL NEURAL NETWORK

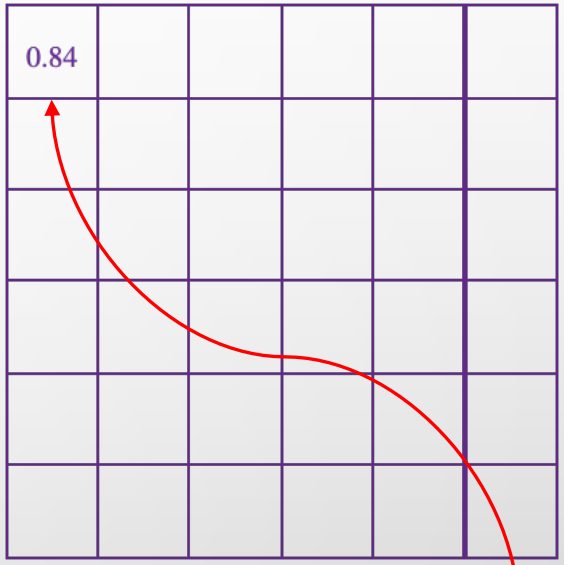
- Convolution
  - Kernel (Filter)

0.37	0.33	−0.76
−0.52	−0.65	0.19
−0.44	−0.03	−0.81

Stride: 1

$b = 0.89$

0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05
−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28
−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45
−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16
−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22
0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0
0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19
0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24



$$z_{11} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.84$$

CONVOLUTIONAL NEURAL NETWORK

- Convolution
  - Kernel (Filter)

0.37	0.33	−0.76
−0.52	−0.65	0.19
−0.44	−0.03	−0.81

Stride: 1

$b = 0.89$

0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05
−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28
−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45
−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16
−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22
0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0
0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19
0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24

0.84	0.88				

$z_{11} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.84$

$z_{12} = \sigma(0.37\times0.18 + 0.33\times0.25 + (-0.76)\times(-0.05) + (-0.52)\times(-0.51) + (-0.65\times-0.65) + 0.19\times(-0.05) + (-0.44\times0.51) + (-0.03\times-0.53) + (-0.81)\times(-0.15) + 0.89) = 0.88$

CONVOLUTIONAL NEURAL NETWORK

- Convolution
  - Kernel (Filter)

0.37	0.33	−0.76
−0.52	−0.65	0.19
−0.44	−0.03	−0.81

Stride: 1

$b = 0.89$

0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05
−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28
−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45
−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16
−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22
0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0
0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19
0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24

0.84	0.88	0.75			

$$z_{11} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.84$$

$$z_{12} = \sigma(0.37\times0.18 + 0.33\times0.25 + (-0.76)\times(-0.05) + (-0.52)\times(-0.51) + (-0.65\times-0.65) + 0.19\times(-0.05) + (-0.44\times0.51) + (-0.03\times-0.53) + (-0.81)\times(-0.15) + 0.89) = 0.88$$

$$z_{13} = \sigma(0.37\times0.25 + 0.33\times(-0.5) + (-0.76)\times0.45 + (-0.52)\times(-0.65) + (-0.65\times-0.05) + 0.19\times(0.29) + (-0.44\times-0.53) + (-0.03\times-0.15) + (-0.81)\times(0.08) + 0.89) = 0.75$$

CONVOLUTIONAL NEURAL NETWORK

- Convolution
  - Kernel (Filter)

0.37	0.33	−0.76
−0.52	−0.65	0.19
−0.44	−0.03	−0.81

Stride: 1

$b = 0.89$

0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05
−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28
−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45
−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16
−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22
0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0
0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19
0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24

0.84	0.88	0.75	0.75		

$$z_{11} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.84$$
$$z_{12} = \sigma(0.37\times0.18 + 0.33\times0.25 + (-0.76)\times(-0.05) + (-0.52)\times(-0.51) + (-0.65\times-0.65) + 0.19\times(-0.05) + (-0.44\times0.51) + (-0.03\times-0.53) + (-0.81)\times(-0.15) + 0.89) = 0.88$$
$$z_{13} = \sigma(0.37\times0.25 + 0.33\times(-0.5) + (-0.76)\times0.45 + (-0.52)\times(-0.65) + (-0.65\times-0.05) + 0.19\times(0.29) + (-0.44\times-0.53) + (-0.03\times-0.15) + (-0.81)\times(0.08) + 0.89) = 0.75$$
$$z_{14} = \sigma(0.37\times(-0.5) + 0.33\times0.45 + (-0.76)\times(-0.04) + (-0.52)\times(-0.05) + (-0.65\times0.29) + 0.19\times0 + (-0.44\times-0.15) + (-0.03\times0.08) + (-0.81)\times(-0.37) + 0.89) = 0.75$$

CONVOLUTIONAL NEURAL NETWORK

- Convolution
  - Kernel (Filter)

0.37	0.33	−0.76
−0.52	−0.65	0.19
−0.44	−0.03	−0.81

Stride: 1

$b = 0.89$

0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05
−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28
−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45
−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16
−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22
0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0
0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19
0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24

0.84	0.88	0.75	0.75	0.84	

$z_{11} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.84$

$z_{12} = \sigma(0.37\times0.18 + 0.33\times0.25 + (-0.76)\times(-0.05) + (-0.52)\times(-0.51) + (-0.65\times-0.65) + 0.19\times(-0.05) + (-0.44\times0.51) + (-0.03\times-0.53) + (-0.81)\times(-0.15) + 0.89) = 0.88$

$z_{13} = \sigma(0.37\times0.25 + 0.33\times(-0.5) + (-0.76)\times0.45 + (-0.52)\times(-0.65) + (-0.65\times-0.05) + 0.19\times(0.29) + (-0.44\times-0.53) + (-0.03\times-0.15) + (-0.81)\times(0.08) + 0.89) = 0.75$

$z_{14} = \sigma(0.37\times(-0.5) + 0.33\times0.45 + (-0.76)\times(-0.04) + (-0.52)\times(-0.05) + (-0.65\times0.29) + 0.19\times0 + (-0.44\times-0.15) + (-0.03\times0.08) + (-0.81)\times(-0.37) + 0.89) = 0.75$

$z_{15} = \sigma(0.37\times0.45 + 0.33\times(-0.04) + (-0.76)\times(-0.73) + (-0.52)\times(0.29) + (-0.65\times0) + 0.19\times(0.06) + (-0.44\times0.08) + (-0.03\times(-0.37)) + (-0.81)\times(-0.25) + 0.89) = 0.84$

CONVOLUTIONAL NEURAL NETWORK

- Convolution
  - Kernel (Filter)

0.37	0.33	−0.76
−0.52	−0.65	0.19
−0.44	−0.03	−0.81

Stride: 1

$b = 0.89$

0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05
−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28
−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45
−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16
−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22
0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0
0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19
0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24

0.84	0.88	0.75	0.75	0.84	0.77

$z_{11} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.84$

$z_{12} = \sigma(0.37\times0.18 + 0.33\times0.25 + (-0.76)\times(-0.05) + (-0.52)\times(-0.51) + (-0.65\times-0.65) + 0.19\times(-0.05) + (-0.44\times0.51) + (-0.03\times-0.53) + (-0.81)\times(-0.15) + 0.89) = 0.88$

$z_{13} = \sigma(0.37\times0.25 + 0.33\times(-0.5) + (-0.76)\times0.45 + (-0.52)\times(-0.65) + (-0.65\times-0.05) + 0.19\times(0.29) + (-0.44\times-0.53) + (-0.03\times-0.15) + (-0.81)\times(0.08) + 0.89) = 0.75$

$z_{14} = \sigma(0.37\times(-0.5) + 0.33\times0.45 + (-0.76)\times(-0.04) + (-0.52)\times(-0.05) + (-0.65\times0.29) + 0.19\times0 + (-0.44\times-0.15) + (-0.03\times0.08) + (-0.81)\times(-0.37) + 0.89) = 0.75$

$z_{15} = \sigma(0.37\times0.45 + 0.33\times(-0.04) + (-0.76)\times(-0.73) + (-0.52)\times(0.29) + (-0.65\times0) + 0.19\times(0.06) + (-0.44\times0.08) + (-0.03\times(-0.37)) + (-0.81)\times(-0.25) + 0.89) = 0.84$

$z_{16} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.77$

CONVOLUTIONAL NEURAL NETWORK

- Convolution
  - Kernel (Filter)

0.37	0.33	−0.76
−0.52	−0.65	0.19
−0.44	−0.03	−0.81

Stride: 1

$b = 0.89$

0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05
−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28
−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45
−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16
−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22
0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0
0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19
0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24

0.84	0.88	0.75	0.75	0.84	0.77
0.79					

$$z_{11} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.84$$
$$z_{12} = \sigma(0.37\times0.18 + 0.33\times0.25 + (-0.76)\times(-0.05) + (-0.52)\times(-0.51) + (-0.65\times-0.65) + 0.19\times(-0.05) + (-0.44\times0.51) + (-0.03\times-0.53) + (-0.81)\times(-0.15) + 0.89) = 0.88$$
$$z_{13} = \sigma(0.37\times0.25 + 0.33\times(-0.5) + (-0.76)\times0.45 + (-0.52)\times(-0.65) + (-0.65\times-0.05) + 0.19\times(0.29) + (-0.44\times-0.53) + (-0.03\times-0.15) + (-0.81)\times(0.08) + 0.89) = 0.75$$
$$z_{14} = \sigma(0.37\times(-0.5) + 0.33\times0.45 + (-0.76)\times(-0.04) + (-0.52)\times(-0.05) + (-0.65\times0.29) + 0.19\times0 + (-0.44\times-0.15) + (-0.03\times0.08) + (-0.81)\times(-0.37) + 0.89) = 0.75$$
$$z_{15} = \sigma(0.37\times0.45 + 0.33\times(-0.04) + (-0.76)\times(-0.73) + (-0.52)\times(0.29) + (-0.65\times0) + 0.19\times(0.06) + (-0.44\times0.08) + (-0.03\times(-0.37)) + (-0.81)\times(-0.25) + 0.89) = 0.84$$
$$z_{16} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.77$$
$$z_{21} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.79$$

CONVOLUTIONAL NEURAL NETWORK

- Convolution
  - Kernel (Filter)

0.37	0.33	−0.76
−0.52	−0.65	0.19
−0.44	−0.03	−0.81

Stride: 1

$b = 0.89$

0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05
−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28
−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45
−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16
−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22
0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0
0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19
0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24

Feature map

0.84	0.88	0.75	0.75	0.84	0.77
0.79	0.64	0.71	0.73	0.79	0.76
0.83	0.75	0.57	0.74	0.71	0.8
0.71	0.46	0.78	0.54	0.76	0.66
0.63	0.78	0.66	0.9	0.79	0.79
0.51	0.54	0.85	0.73	0.78	0.82

$$z_{11} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.84$$
$$z_{12} = \sigma(0.37\times0.18 + 0.33\times0.25 + (-0.76)\times(-0.05) + (-0.52)\times(-0.51) + (-0.65\times-0.65) + 0.19\times(-0.05) + (-0.44\times0.51) + (-0.03\times-0.53) + (-0.81)\times(-0.15) + 0.89) = 0.88$$
$$z_{13} = \sigma(0.37\times0.25 + 0.33\times(-0.5) + (-0.76)\times0.45 + (-0.52)\times(-0.65) + (-0.65\times-0.05) + 0.19\times(0.29) + (-0.44\times-0.53) + (-0.03\times-0.15) + (-0.81)\times(0.08) + 0.89) = 0.75$$
$$z_{14} = \sigma(0.37\times(-0.5) + 0.33\times0.45 + (-0.76)\times(-0.04) + (-0.52)\times(-0.05) + (-0.65\times0.29) + 0.19\times0 + (-0.44\times-0.15) + (-0.03\times0.08) + (-0.81)\times(-0.37) + 0.89) = 0.75$$
$$z_{15} = \sigma(0.37\times0.45 + 0.33\times(-0.04) + (-0.76)\times(-0.73) + (-0.52)\times(0.29) + (-0.65\times0) + 0.19\times(0.06) + (-0.44\times0.08) + (-0.03\times(-0.37)) + (-0.81)\times(-0.25) + 0.89) = 0.84$$
$$z_{16} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.77$$
$$z_{21} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.79$$
$$\vdots$$
$$z_{66} = \sigma(0.37\times0.37 + 0.33\times0.11 + (-0.76)\times0 + (-0.52)\times(-0.89) + (-0.65\times-0.06) + 0.19\times(-0.19) + (-0.44\times-0.49) + (-0.03\times0.04) + (-0.81)\times0.24 + 0.89) = 0.82$$

CONVOLUTIONAL NEURAL NETWORK

- Convolution
  - Kernel (Filter)

0.37	0.33	−0.76
−0.52	−0.65	0.19
−0.44	−0.03	−0.81

$b = 0.89$

0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05
−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28
−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45
−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16
−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22
0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0
0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19
0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24

What happens when the stride is 2?

CONVOLUTIONAL NEURAL NETWORK

- Convolution
  - Kernel (Filter)

0.37	0.33	−0.76
−0.52	−0.65	0.19
−0.44	−0.03	−0.81

Stride: 2

$b = 0.89$

0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05
−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28
−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45
−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16
−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22
0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0
0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19
0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24

0.84		

$$z_{11} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.84$$

CONVOLUTIONAL NEURAL NETWORK

- Convolution
  - Kernel (Filter)

0.37	0.33	−0.76
−0.52	−0.65	0.19
−0.44	−0.03	−0.81

Stride: 2

$b = 0.89$

0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05
−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28
−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45
−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16
−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22
0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0
0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19
0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24

0.84	0.75	

$$z_{11} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.84$$

$$z_{12} = \sigma(0.37\times0.25 + 0.33\times(-0.5) + (-0.76)\times0.45 + (-0.52)\times(-0.65) + (-0.65\times-0.05) + 0.19\times(0.29) + (-0.44\times-0.53) + (-0.03\times-0.15) + (-0.81)\times(0.08) + 0.89) = 0.75$$

CONVOLUTIONAL NEURAL NETWORK

- Convolution
  - Kernel (Filter)

0.37	0.33	−0.76
−0.52	−0.65	0.19
−0.44	−0.03	−0.81

Stride: 2

$b = 0.89$

0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05
−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28
−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45
−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16
−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22
0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0
0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19
0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24

0.84	0.75	0.84

$$z_{11} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.84$$

$$z_{12} = \sigma(0.37\times0.25 + 0.33\times(-0.5) + (-0.76)\times0.45 + (-0.52)\times(-0.65) + (-0.65\times-0.05) + 0.19\times(0.29) + (-0.44\times-0.53) + (-0.03\times-0.15) + (-0.81)\times(0.08) + 0.89) = 0.75$$

$$z_{13} = \sigma(0.37\times0.45 + 0.33\times(-0.04) + (-0.76)\times(-0.73) + (-0.52)\times(0.29) + (-0.65\times0) + 0.19\times(0.06) + (-0.44\times0.08) + (-0.03\times(-0.37)) + (-0.81)\times(-0.25) + 0.89) = 0.84$$

CONVOLUTIONAL NEURAL NETWORK

- Convolution
  - Kernel (Filter)

0.37	0.33	−0.76
−0.52	−0.65	0.19
−0.44	−0.03	−0.81

Stride: 2

$b = 0.89$

0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05
−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28
−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45
−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16
−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22
0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0
0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19
0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24

0.84	0.75	0.84
0.83		

$$z_{11} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.84$$

$$z_{12} = \sigma(0.37\times0.25 + 0.33\times(-0.5) + (-0.76)\times0.45 + (-0.52)\times(-0.65) + (-0.65\times-0.05) + 0.19\times(0.29) + (-0.44\times-0.53) + (-0.03\times-0.15) + (-0.81)\times(0.08) + 0.89) = 0.75$$

$$z_{13} = \sigma(0.37\times0.45 + 0.33\times(-0.04) + (-0.76)\times(-0.73) + (-0.52)\times(0.29) + (-0.65\times0) + 0.19\times(0.06) + (-0.44\times0.08) + (-0.03\times(-0.37)) + (-0.81)\times(-0.25) + 0.89) = 0.84$$

$$z_{21} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.79$$

CONVOLUTIONAL NEURAL NETWORK

- Convolution
  - Kernel (Filter)

0.37	0.33	−0.76
−0.52	−0.65	0.19
−0.44	−0.03	−0.81

Stride: 2

$b = 0.89$

0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05
−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28
−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45
−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16
−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22
0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0
0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19
0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24

Feature map

0.84	0.75	0.84
0.83	0.57	0.71
0.63	0.66	0.79

$$z_{11} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.84$$

$$z_{12} = \sigma(0.37\times0.25 + 0.33\times(-0.5) + (-0.76)\times0.45 + (-0.52)\times(-0.65) + (-0.65\times-0.05) + 0.19\times(0.29) + (-0.44\times-0.53) + (-0.03\times-0.15) + (-0.81)\times(0.08) + 0.89) = 0.75$$

$$z_{13} = \sigma(0.37\times0.45 + 0.33\times(-0.04) + (-0.76)\times(-0.73) + (-0.52)\times(0.29) + (-0.65\times0) + 0.19\times(0.06) + (-0.44\times0.08) + (-0.03\times(-0.37)) + (-0.81)\times(-0.25) + 0.89) = 0.84$$

$$z_{21} = \sigma(0.37\times0.15 + 0.33\times0.18 + (-0.76)\times0.25 + (-0.52)\times(-0.25) + (-0.65\times-0.51) + 0.19\times(-0.65) + (-0.44\times-0.24) + (-0.03\times0.51) + (-0.81)\times(-0.53) + 0.89) = 0.79$$

$$\vdots$$

$$z_{33} = \sigma(0.37\times0.37 + 0.33\times0.11 + (-0.76)\times0 + (-0.52)\times(-0.89) + (-0.65\times-0.06) + 0.19\times(-0.19) + (-0.44\times-0.49) + (-0.03\times0.04) + (-0.81)\times0.24 + 0.89) = 0.82$$

PADDING

- Padding is used to preserve the dimensions of input image in the feature map
  - The input image before convolution and the feature map after convolution would have the same size
- Types of padding:
  - Same padding
  - Zero padding
- Padding with even-sized kernels
  - Non-symmetric padding needs to be used for even-sized kernels
  - This is one reason why odd-sized kernels are preferred

PADDING

0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05
−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28
−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45
−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16
−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22
0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0
0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19
0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24

Zero padding

0	0	0	0	0	0	0	0	0	0
0	0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05	0
0	−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28	0
0	−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45	0
0	−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16	0
0	−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22	0
0	0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0	0
0	0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19	0
0	0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24	0
0	0	0	0	0	0	0	0	0	0

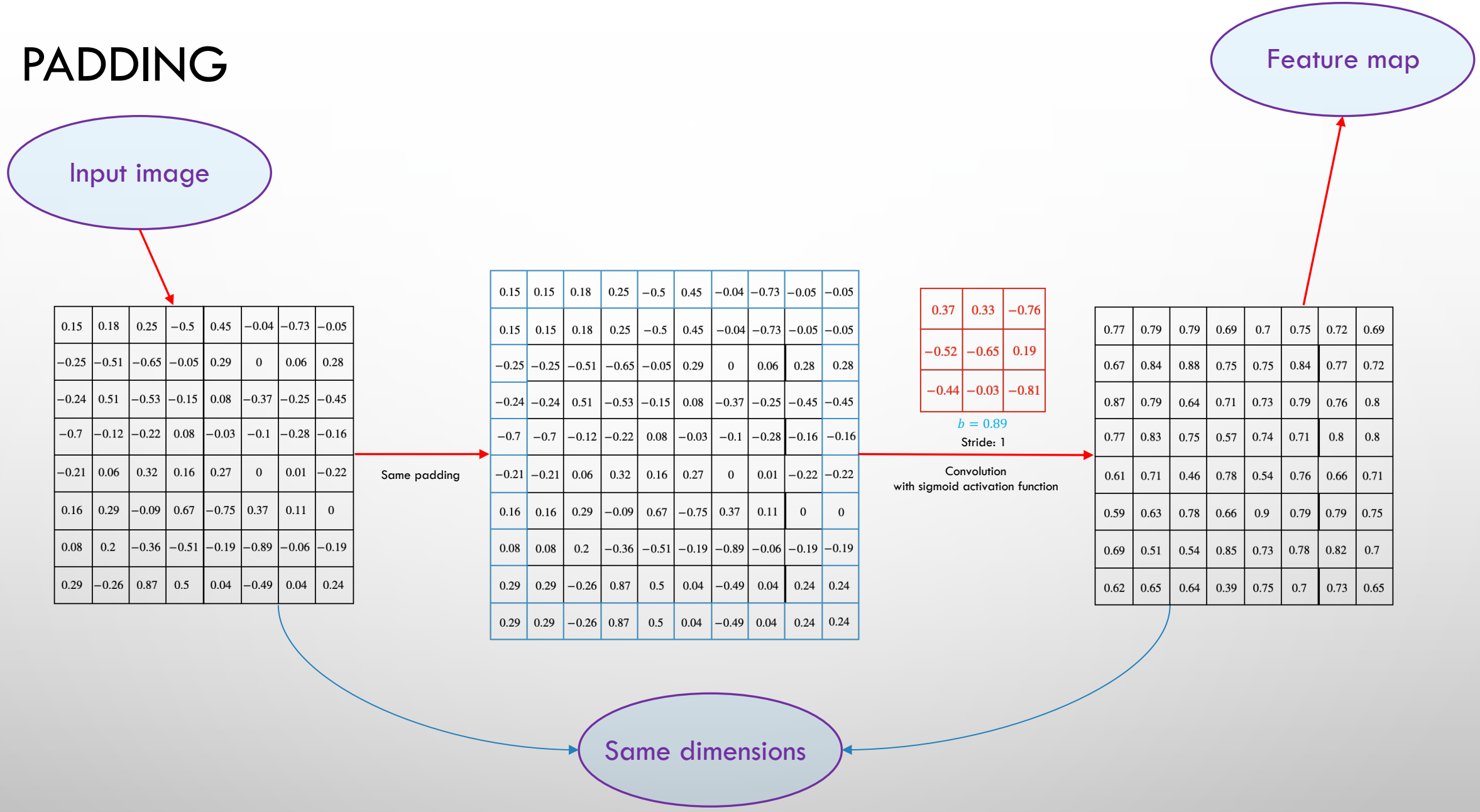
PADDING

0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05
−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28
−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45
−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16
−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22
0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0
0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19
0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24

Same padding

0.15	0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05	−0.05
0.15	0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05	−0.05
−0.25	−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28	0.28
−0.24	−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45	−0.45
−0.7	−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16	−0.16
−0.21	−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22	−0.22
0.16	0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0	0
0.08	0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19	−0.19
0.29	0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24	0.24
0.29	0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24	0.24

PADDING


Input image
Feature map
0.15 0.18 0.25 −0.5 0.45 −0.04 −0.73 −0.05
−0.25 −0.51 −0.65 −0.05 0.29 0 0.06 0.28
−0.24 0.51 −0.53 −0.15 0.08 −0.37 −0.25 −0.45
−0.7 −0.12 −0.22 0.08 −0.03 −0.1 −0.28 −0.16
−0.21 0.06 0.32 0.16 0.27 0 0.01 −0.22
0.16 0.29 −0.09 0.67 −0.75 0.37 0.11 0
0.08 0.2 −0.36 −0.51 −0.19 −0.89 −0.06 −0.19
0.29 −0.26 0.87 0.5 0.04 −0.49 0.04 0.24
Same padding
0.15 0.15 0.18 0.25 −0.5 0.45 −0.04 −0.73 −0.05 −0.05
0.15 0.15 0.18 0.25 −0.5 0.45 −0.04 −0.73 −0.05 −0.05
−0.25 −0.25 −0.51 −0.65 −0.05 0.29 0 0.06 0.28 0.28
−0.24 −0.24 0.51 −0.53 −0.15 0.08 −0.37 −0.25 −0.45 −0.45
−0.7 −0.7 −0.12 −0.22 0.08 −0.03 −0.1 −0.28 −0.16 −0.16
−0.21 −0.21 0.06 0.32 0.16 0.27 0 0.01 −0.22 −0.22
0.16 0.16 0.29 −0.09 0.67 −0.75 0.37 0.11 0 0
0.08 0.08 0.2 −0.36 −0.51 −0.19 −0.89 −0.06 −0.19 −0.19
0.29 0.29 −0.26 0.87 0.5 0.04 −0.49 0.04 0.24 0.24
0.29 0.29 −0.26 0.87 0.5 0.04 −0.49 0.04 0.24 0.24
0.37 0.33 −0.76
−0.52 −0.65 0.19
−0.44 −0.03 −0.81
b = 0.89
Stride: 1
Convolution
with sigmoid activation function
0.77 0.79 0.79 0.69 0.7 0.75 0.72 0.69
0.67 0.84 0.88 0.75 0.75 0.84 0.77 0.72
0.87 0.79 0.64 0.71 0.73 0.79 0.76 0.8
0.77 0.83 0.75 0.57 0.74 0.71 0.8 0.8
0.61 0.71 0.46 0.78 0.54 0.76 0.66 0.71
0.59 0.63 0.78 0.66 0.9 0.79 0.79 0.75
0.69 0.51 0.54 0.85 0.73 0.78 0.82 0.7
0.62 0.65 0.64 0.39 0.75 0.7 0.73 0.65
Same dimensions

MAX POOLING

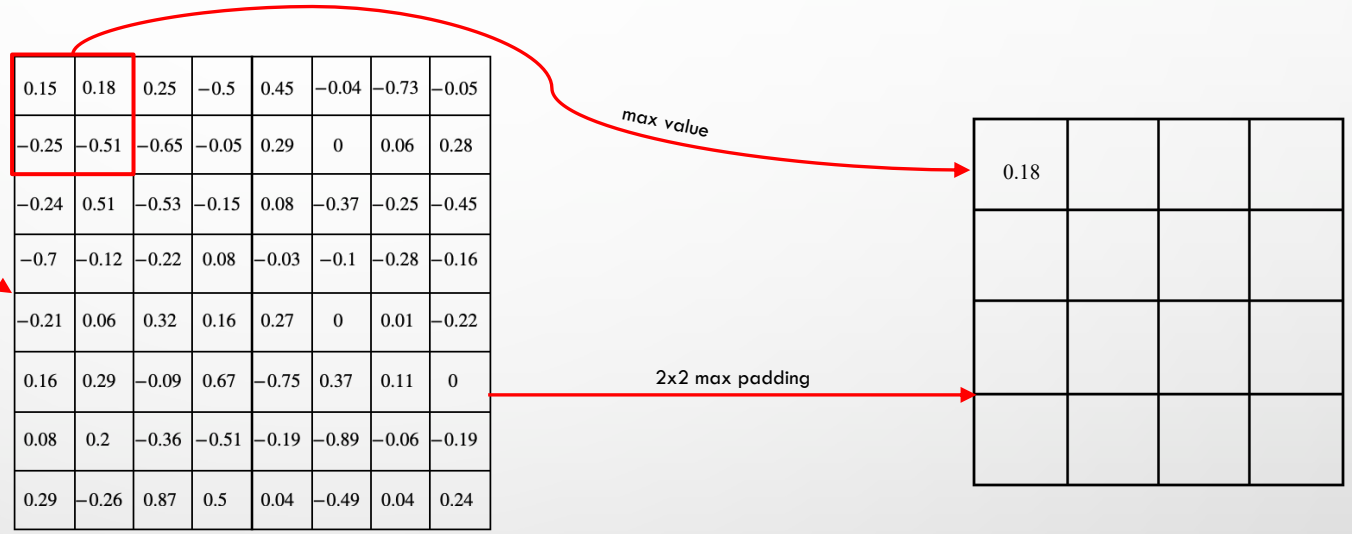
- Pooling is used to reduce the spatial dimensions, which can help in
  - Reduction in the number of parameters
  - Being less prone to overfitting
- Commonly used pooling types:
  - Max pooling
  - Avg pooling

MAX POOLING

Feature map before pooling

2x2 max pooling:

The maximum value of each 2x2 block is computed

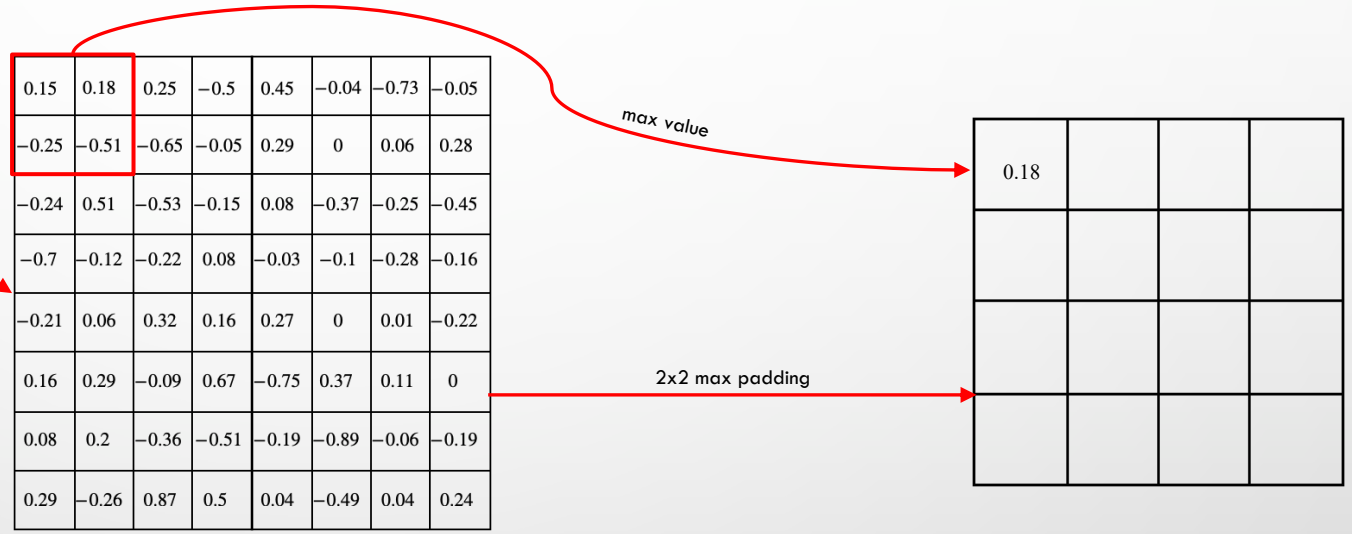


MAX POOLING

Feature map before pooling

2x2 max pooling:

The maximum value of each 2x2 block is computed

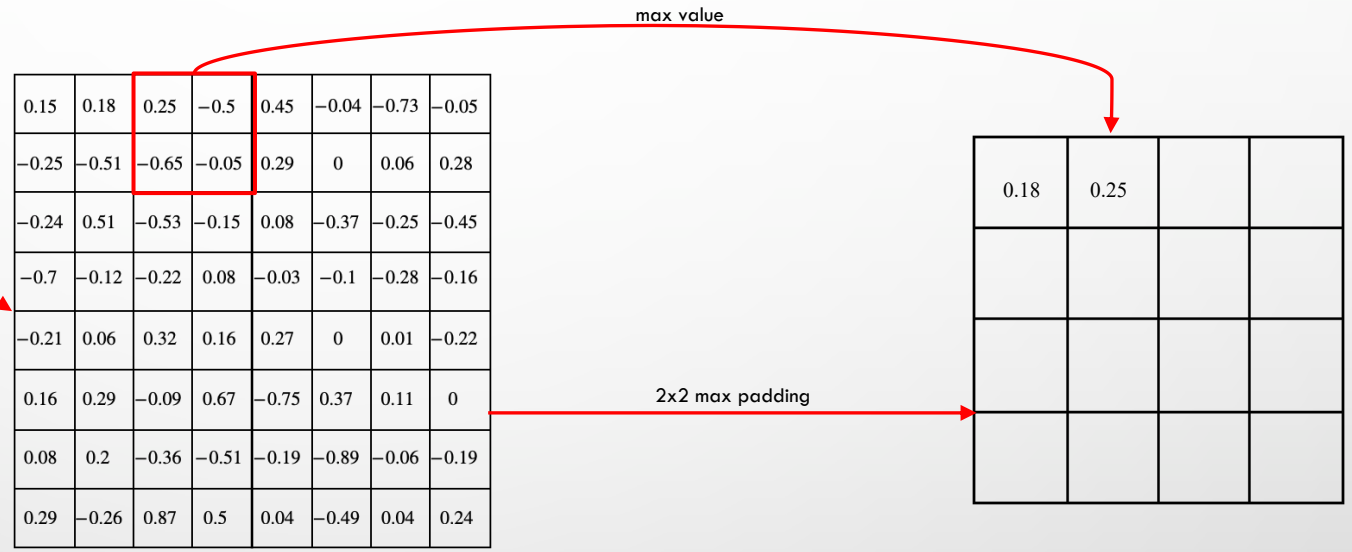


MAX POOLING

Feature map before pooling

2x2 max pooling:

The maximum value of each 2x2 block is computed

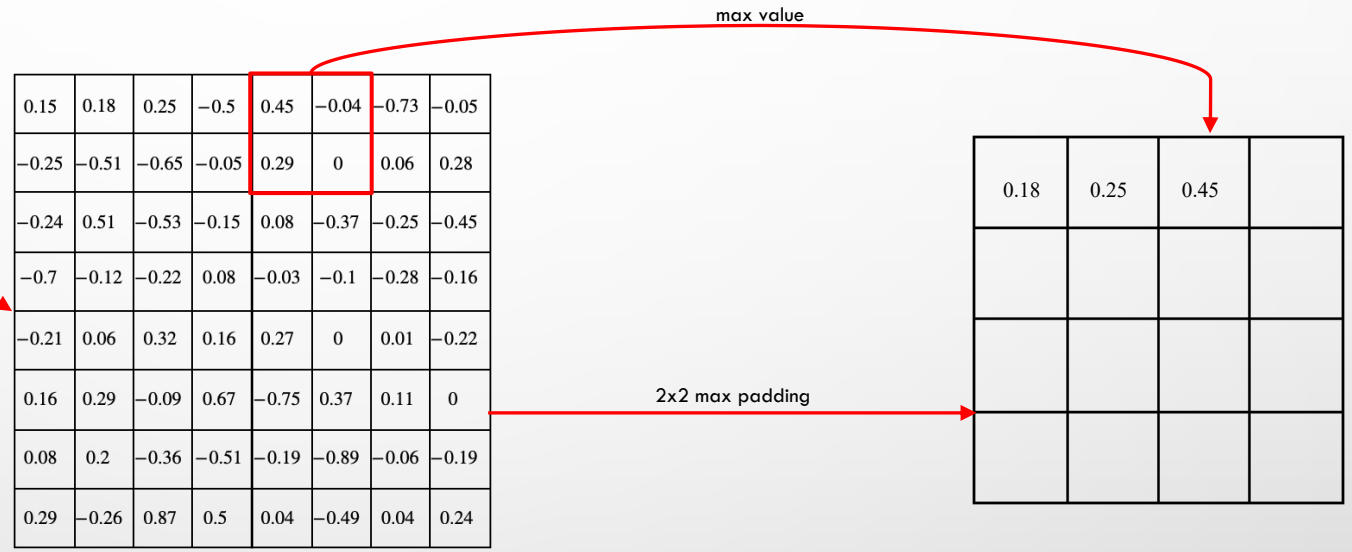


MAX POOLING

Feature map before pooling

2x2 max pooling:

The maximum value of each 2x2 block is computed

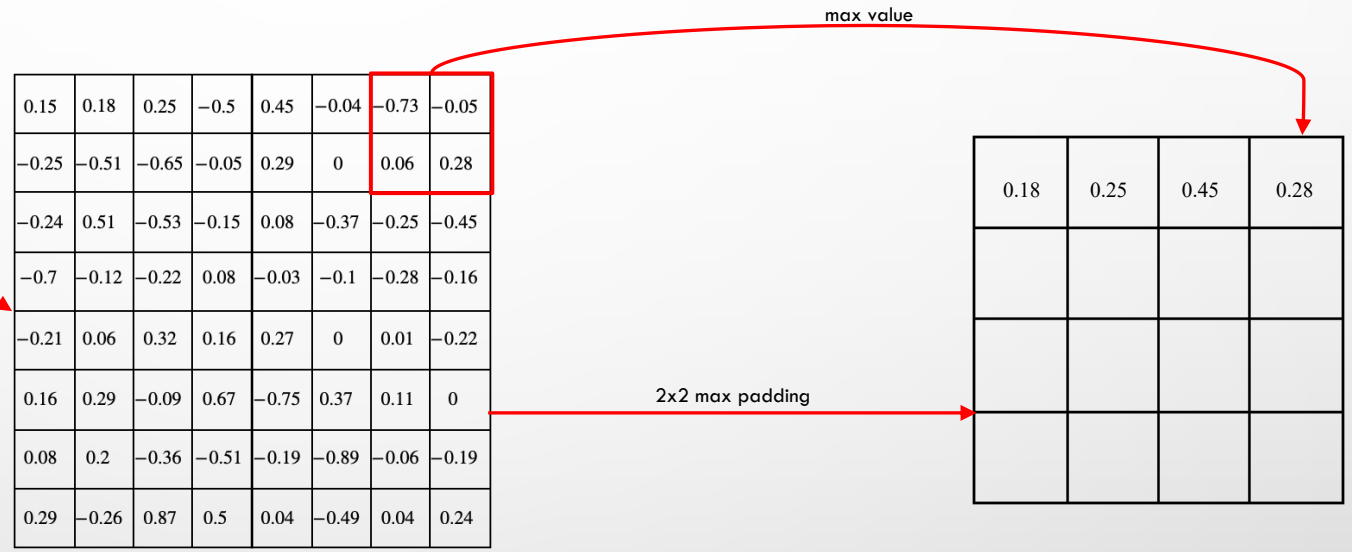


MAX POOLING

Feature map before pooling

2x2 max pooling:

The maximum value of each 2x2 block is computed

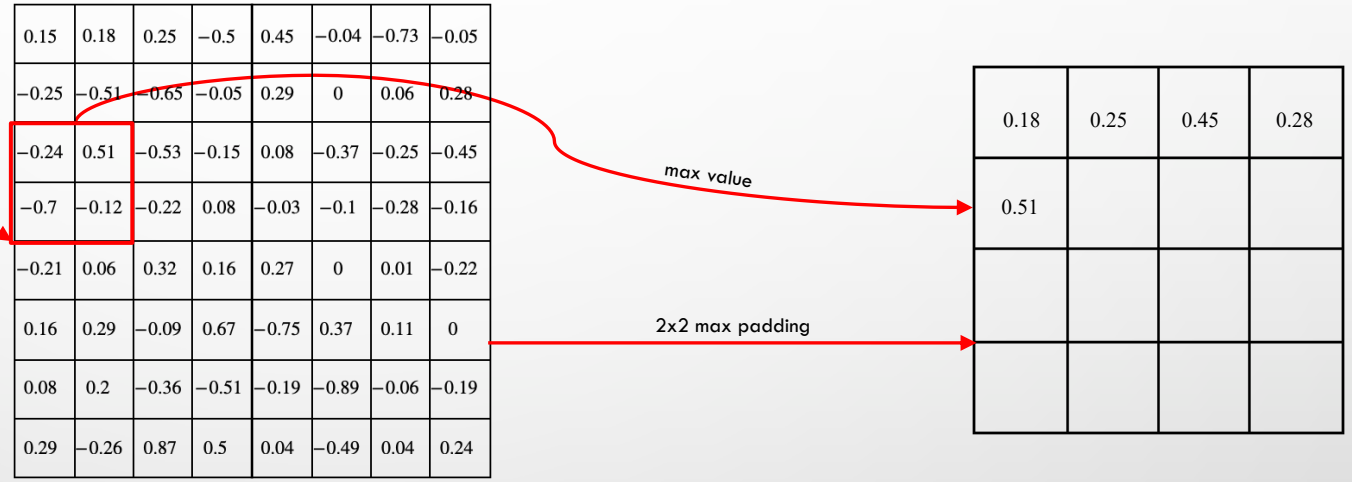


MAX POOLING

Feature map
before pooling

2x2 max pooling:

The maximum value of each 2x2 block is computed

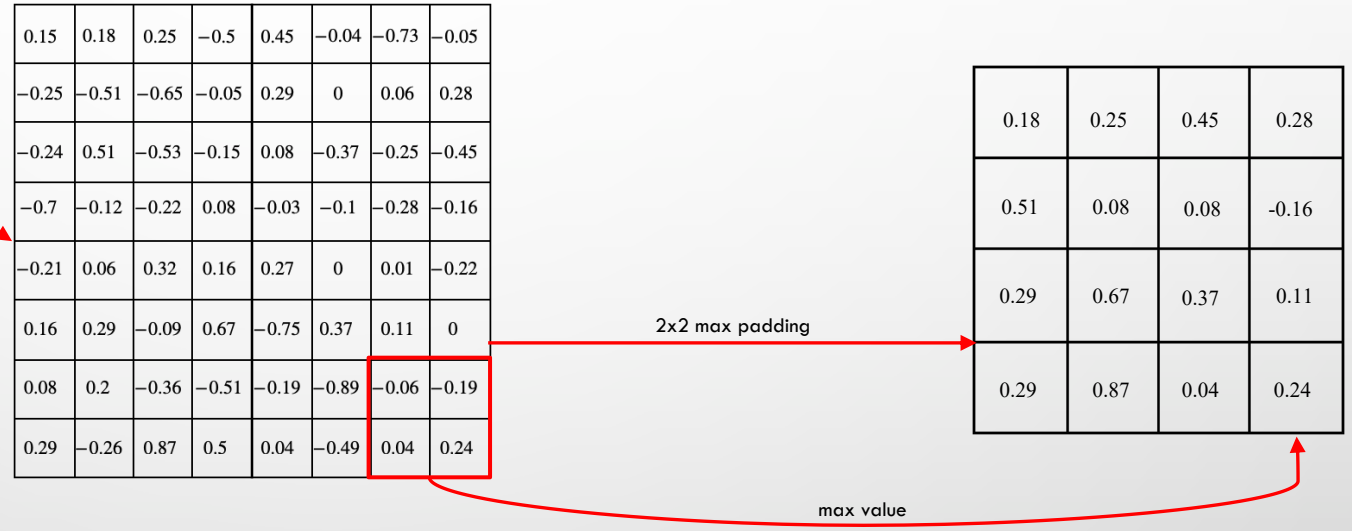


MAX POOLING

Feature map before pooling

2x2 max pooling:

The maximum value of each 2x2 block is computed



MAX POOLING

Feature map before pooling

2x2 max pooling:

The maximum value of each 2x2 block is computed

0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05
−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28
−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45
−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16
−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22
0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0
0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19
0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24

2x2 max padding

0.18	0.25	0.45	0.28
0.51	0.08	0.08	-0.16
0.29	0.67	0.37	0.11
0.29	0.87	0.04	0.24

Feature map after pooling

MAX POOLING

Feature map before pooling

Avg pooling:

In average pooling the average value is computed instead of the maximum value

0.15	0.18	0.25	−0.5	0.45	−0.04	−0.73	−0.05
−0.25	−0.51	−0.65	−0.05	0.29	0	0.06	0.28
−0.24	0.51	−0.53	−0.15	0.08	−0.37	−0.25	−0.45
−0.7	−0.12	−0.22	0.08	−0.03	−0.1	−0.28	−0.16
−0.21	0.06	0.32	0.16	0.27	0	0.01	−0.22
0.16	0.29	−0.09	0.67	−0.75	0.37	0.11	0
0.08	0.2	−0.36	−0.51	−0.19	−0.89	−0.06	−0.19
0.29	−0.26	0.87	0.5	0.04	−0.49	0.04	0.24

2x2 avg padding

-0.11	-0.24	0.18	-0.11
-0.14	-0.2	-0.1	-0.28
0.08	0.26	-0.03	-0.02
0.08	0.12	-0.38	0.01

Feature map after pooling